



# Phantoms for evaluating the impact of skin pigmentation on photoacoustic imaging and oximetry performance

WILLIAM C. VOGT,<sup>\*</sup>  KEITH A. WEAR,<sup>\*</sup>  AND T. JOSHUA PFEFER 

Center for Devices and Radiological Health, U.S. Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD 20993, USA

<sup>\*</sup>[william.vogt@fda.hhs.gov](mailto:william.vogt@fda.hhs.gov)

**Abstract:** Recent reports have raised concerns of potential racial disparities in performance of optical oximetry technologies. To investigate how variable epidermal melanin content affects performance of photoacoustic imaging (PAI) devices, we developed plastisol phantoms combining swappable skin-mimicking layers with a breast phantom containing either India ink or blood adjusted to 50–100% SO<sub>2</sub> using sodium dithionite. Increasing skin pigmentation decreased maximum imaging depth by up to 25%, enhanced image clutter, and increased root-mean-square error in SO<sub>2</sub> from 8.0 to 17.6% due to signal attenuation and spectral coloring effects. This phantom tool can aid in evaluating PAI device robustness to ensure high performance in all patients.

## 1. Introduction

Clinical study results have raised concerns regarding the potential for racial disparities in optical diagnostic device performance, particularly of oximetry devices. Most notably, several studies have reported disparities in pulse oximeter readings between patients of different self-identified race/ethnicity [1–3]. One potential confounder thought to contribute to these reported effects is skin melanin content or pigmentation level, which varies widely in humans [4]. As most optical diagnostic technologies involve dermal or trans-dermal measurements, differences in epidermal melanin content can lead to alterations in light transport effects and detected signals. Several studies have investigated the impact of skin pigmentation on various optical devices including near-infrared spectroscopy (NIRS)-based cerebral oximeters [5,6], bilirubinometers [7,8], and hyperspectral reflectance imagers [9].

Photoacoustic imaging (PAI) enables deep mapping of tissue vasculature and blood oxygen saturation (SO<sub>2</sub>) for clinical applications including cancer detection (especially breast cancer), oximetry, and surgical guidance [10–13]. Melanin is a highly absorptive chromophore that can be detected by PAI, and several studies have proposed melanin as a target of interest in clinical PAI applications including melanoma detection [14], lesion depth assessment [15], circulating melanoma cell detection [16], or the use of melanin-derived contrast agents [17]. However, recent studies have begun to provide clinical evidence that photoacoustic images and oximetry signals can be adversely affected by epidermal melanin as a confounding factor. Li et al., reported increased brightness and clutter with increasing skin pigmentation level in human subjects imaged using a photoacoustic mesoscopy system [18], and Mantri and Jokerst observed that PAI-based measurements of SO<sub>2</sub> in the radial artery in human subjects decreased with increasing Fitzpatrick skin phototype [19]. These findings demonstrate the need to better understand and address skin pigmentation as a factor impacting photoacoustic image quality and oximetry measurement accuracy.

Medical devices should be tested during development to ensure acceptable performance in all patients, including across patients with different skin pigmentation levels. Clinical testing of optical oximetry device performance, for instance through controlled oxygen desaturation studies in healthy adult volunteers [20], carries significant costs, poses nonnegligible risks

to human subjects, and cannot achieve a full range of  $\text{SO}_2$  values due to safety limits of desaturation. The availability of suitable bench test methods, such as those based on tissue-mimicking phantoms, would enable cheaper, faster assessment of device performance and pigmentation robustness earlier in the device development process. Such methods can also facilitate quality assurance measurements and device comparison. While there have been significant advancements in development of tissue phantoms for PAI [21], there remains an outstanding need for well-characterized phantoms capable of assessing PAI system robustness to variable skin pigmentation.

Skin-mimicking phantoms have been developed for other optical diagnostic technologies. Saager et al. developed thin polydimethylsiloxane (PDMS) layers with adjustable optical properties [22], and our group previously developed PDMS skin-simulating layers that can be combined with a 3D printed cerebrovascular module to evaluate NIRS cerebral oximeter performance for different skin pigmentation levels [6]. However, currently available skin-mimicking materials and phantoms are not suitable for PAI. PDMS does not have biologically relevant acoustic properties; while soft tissue has a speed of sound from  $\sim 1450$ – $1575$  m/s and acoustic attenuation from  $\sim 0.5$ – $30$  dB/cm over 1–10 MHz, one study reported PDMS formulations having lower speed of sound (1076–1119 m/s) and higher acoustic attenuation ( $\sim 10$ – $65$  dB/cm over 3–11 MHz) [23]. These high attenuation values would cause phantoms to produce weak signals and underestimate device sensitivity and imaging depth. Similarly, our preliminary evaluation of some 3D printing resins suggested they may not be suitable for ultrasound or photoacoustic imaging due to excessive acoustic attenuation, high speed of sound, and strong image artifacts (unpublished data).

Our overall goal is to advance racial equity in healthcare by supporting development and evaluation of robust PAI medical devices. There is an outstanding need for tools that can characterize PAI device susceptibility to variable skin pigmentation and support testing of device design improvements intended to compensate for these potential effects. To this end, our study objectives were to develop tissue-mimicking phantoms that simulate the range of near-infrared skin absorption due to melanin and to develop phantom-based test methods that enable evaluation of PAI device image quality and oximetry measurement accuracy across a biologically relevant range of epidermal melanin concentrations.

## 2. Methods

### 2.1. Material development and characterization

We opted to develop skin tissue-mimicking materials (TMMs) based on our previously developed polyvinyl chloride plastisol (PVCP) phantom material [24]. Our goal was to develop a formulation with melanin-mimicking absorption; since melanin was difficult to solubilize in PVCP, we searched for absorptive dyes that could replicate melanin absorption in the near-infrared range. We screened several dyes for PVCP compatibility and desirable optical properties, then selected alcohol-soluble nigrosin (Sigma-Aldrich, St. Louis, MO) for further development. Nigrosin-ethanol solutions with concentrations of 0.125–1.50 mg/mL were also imaged using our custom PAI system to verify production of strong photoacoustic signals and melanin-like spectral response. PVCP optical scattering was adjusted using 0–5 mg/mL anatase titanium dioxide ( $\text{TiO}_2$ , Sigma-Aldrich). Nigrosin was added to PVCP at 0–2 mg/mL, either by dissolving into di(2-ethylhexyl) adipate (DEHA, Sigma-Aldrich) or as a concentrated solution in 99% ethanol that was added to PVCP at 2% by volume. As our previous work showed PVCP properties can be tuned by varying the ratio of binary plasticizer mixtures, we also added nigrosin to liquid PVCP formulations with varying composition of DEHA and benzyl butyl phthalate (BBP, Sigma-Aldrich) from 0–100% BBP in 25% steps. All PVCP formulations contained 1% heat stabilizer (M-F Manufacturing, Inc., Fort Worth, TX).

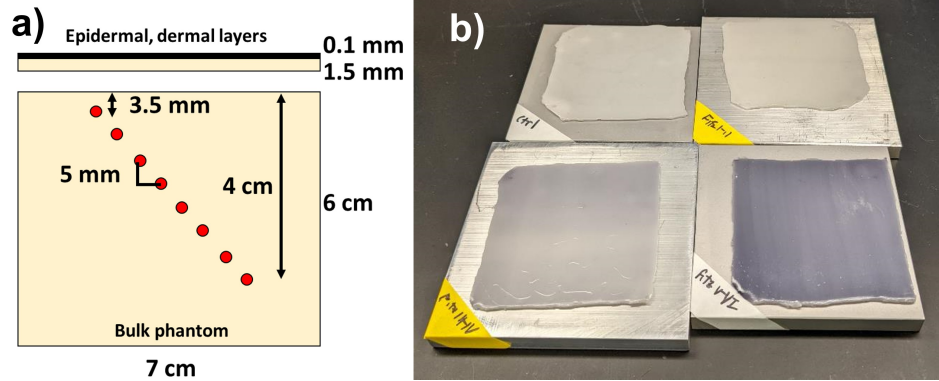
For each formulation, PVCP was prepared by heating within a round bottom flask suspended within an oil bath as described previously [24]. PVCP samples with dimensions 38 mm x 38 mm x 0.5 mm or 1.0 mm were molded for characterization by integrating sphere spectrophotometry over 400–1100 nm. To ensure adequate signal-to-noise ratio in characterization samples used for spectrophotometry, samples with low nigrosin concentrations (0.0625, 0.125, and 0.25 mg/mL) had a thickness of 1 mm, while higher concentration samples had a thickness of 0.5 mm. Samples were prepared using a two-face aluminum mold, where spacers made of aluminum shim stock were glued to one plate to impose the appropriate sample thickness. Sample thickness was measured using calipers. Non-turbid samples were scanned using a traditional spectrophotometry configuration, and extinction coefficient was computed from measured transmittance following Beer's Law. Turbid PVCP samples (i.e., those containing TiO<sub>2</sub>) were scanned using an integrating sphere configuration, with optical absorption and reduced scattering coefficients calculated from total transmittance and diffuse reflectance measurements using the inverse adding-doubling method [25]. Acoustic characterization of 38 mm-diameter, 5-mm-thick disk samples with 15% m/m PVC resin, 3:1 BBP:DEHA, and with or without 2% v/v ethanol was performed from 6–18 MHz as previously described to determine the impact of new additives on speed of sound and acoustic attenuation [26].

Available melanin extinction spectra in the literature are available for comparison against the developed phantoms, but typically do not extend into the near-infrared beyond 800 nm [27–29]. To facilitate comparison of nigrosin-doped PVCP extinction spectra to that of melanin, spectrophotometry was also performed on synthetic, tyrosine-derived melanin powder (M0418, Sigma Aldrich), which was dissolved in 0.01 M NaOH. Melanin solutions were pipetted into 1 mm-thick quartz cuvettes and scanned using traditional spectrophotometry as above.

## 2.2. Phantom construction

We elected to develop a modular phantom approach comprised of a base bulk phantom containing fillable tubes at various depths and a set of separate skin-mimicking layers (Fig. 1(a)). The bulk PVCP phantom contained an array of K-flex PVDF tubes (Eldon James, Fort Collins, CO) with 0.8 mm inner diameter. This tubing material was selected for its low oxygen gas permeability, as initial testing of various tube materials indicated that static blood volumes at low SO<sub>2</sub> injected into phantom tubes can quickly reoxygenate over time, even with the tube ends sealed (see Supplement 1). These data suggested that oxygen gas transfer through the tube wall and PVCP was substantial, with reoxygenation occurring within a few minutes, but K-flex afforded necessary oxygen stability (Fig. S1). K-flex tubing is also translucent and colorless by visual inspection, which may indicate minimal spectral coloring in the tube wall compared to other impermeable tubing options.. This tubing does not exhibit significant PAI signal over 700–900 nm (Fig. S2).

Melanin is produced by melanocytes in the basal layer of the epidermis and packaged into melanosomes, densely packed melanin granules that then migrate through the epidermis [30]. Thus, skin-mimicking phantoms intended to replicate melanin absorption should possess a very thin, highly absorptive epidermal layer. Breast imaging is a major clinical application of PAI, and the thicknesses of epidermis and dermis in female breast have been reported as  $76.9 \pm 26.2$   $\mu\text{m}$  [31] and  $1.45 \pm 0.3$  mm, respectively [32]. From initial testing, PVCP films fabricated with epidermis-like thickness lacked sufficient mechanical strength to enable use as a separate phantom layer. Thus, we fabricated two-layer skin phantoms consisting of a thin ( $< 100$   $\mu\text{m}$ ) epidermal layer bonded to a 1.5 mm-thick dermis-mimicking layer (Fig. 1(a)). Aluminum well molds were machined with a 7 cm x 7 cm x 100  $\mu\text{m}$ -deep well, onto which 15 mL of epidermis-mimicking PVCP was poured. After pouring, the PVCP was immediately planed flat using the edge of an aluminum plate. After cooling for 30 minutes, a batch of dermis-mimicking PVCP was prepared and poured on top of the epidermis layer and immediately compressed with an aluminum plate possessing 1.5 mm spacers to control final thickness of the combined epidermis-dermis phantom.



**Fig. 1.** a) PVCP phantom schematic showing removal/swappable skin layer component. b) Photograph of set of skin-mimicking layers representing control (upper left), light (upper right), moderate (lower left), and dark (lower right) layers.

After cooling for 1 hour, the plates were separated and excess PVCP was trimmed away, leaving a two-component bonded skin layer that was durable enough for handling. To measure typical epidermis layer thickness, a PVCP skin phantom was created with a turbid epidermal layer containing 5 mg/mL  $\text{TiO}_2$  and a clear dermis layer without  $\text{TiO}_2$ . This design produced high contrast when imaged using a custom 1070 nm spectral-domain optical coherence tomography (OCT) system with 6  $\mu\text{m}$  axial resolution, 25  $\mu\text{m}$  lateral resolution (see Supplement 1) [33].

We fabricated a set of four skin-mimicking phantom layers (control, light, moderate, and dark layers, Fig. 1(b)) to conduct a parametric study of skin pigmentation effects on photoacoustic image quality and oximetry measurement accuracy. There are limited data mapping skin optical absorption coefficient to melanin concentration, but estimates from Jacques suggest melanosome volume fraction ( $M_f$ ) values for light, moderate, and dark skin of 1.3-6.3%, 11-16%, and 18-43%, respectively [34]. These values are a descriptive convention, rather than firm specifications. We thus chose target  $M_f$  values of 0% for the control layer, 1% for light skin, 15% for moderate skin, and 43% for dark skin. Literature data for the extinction spectrum of melanosomes were used to convert these  $M_f$  values into corresponding  $\mu_a$  values at 800 nm following Jacques' data for melanosome absorption coefficient,  $\mu_{a,\text{melanosome}}$  [29]:

$$\mu_a(\lambda) = M_f \mu_{a,\text{melanosome}}(\lambda) = M_f (519 \text{ cm}^{-1}) \left( \frac{\lambda}{500 \text{ nm}} \right)^{-3.53}$$

This yields target  $\mu_a$  values at 800 nm of 0, 1, 16, and 42  $\text{cm}^{-1}$  for the control, light, moderate, and dark layers, respectively. The control layer absorption is nonzero ( $\sim 0.1 \text{ cm}^{-1}$ ) due to light absorption by PVCP. The scattering coefficient,  $\mu_s'$ , for the epidermal layer for all cases was chosen to be 20  $\text{cm}^{-1}$  at 800 nm based on measurements reported in Asian epidermis samples [35].

The bulk phantom was made using our previously described breast-mimicking PVCP formulation, using a 3:1 ratio of BBP/DEHA, 10% m/m PVC resin, 1% heat stabilizer, and 2 mg/mL  $\text{TiO}_2$  [24,36]. This formulation yields nominal values of  $\mu_a = \sim 0.1 \text{ cm}^{-1}$  and  $\mu_s' = \sim 10 \text{ cm}^{-1}$  at 800 nm, as well as a speed of sound of 1467 m/s and acoustic attenuation of 0.22  $f^{1.76}$  dB/cm for frequencies,  $f$ , from 4–9 MHz. To isolate the effects of variable epidermal optical properties on photoacoustic images, the dermis-mimicking layer was comprised of the same formulation as the bulk phantom. This would reduce acoustic impedance mismatch at the dermis/bulk interface, minimizing acoustic artifacts due to specular reflection.



### 2.3. Imaging and data analysis

Two imaging experiments were performed. First, to test impact of skin pigmentation on spectral signal content and overall image quality, the phantom was imaged with its tubes filled with India ink solutions with  $\mu_a$  of  $4 \text{ cm}^{-1}$  at 800 nm (near the isosbestic point of hemoglobin), corresponding to a total hemoglobin concentration (tHb) of  $\sim 13.5 \text{ g/dL}$ . For oximetry testing, phantom tubes were injected with whole, defibrinated bovine blood (Lampire Biological Laboratories, Pipersville, PA). Blood  $\text{SO}_2$  was adjusted by adding sodium dithionite. For each  $\text{SO}_2$  setpoint, a 10 mL whole blood sample was prepared in a 20 mL vial, to which 1 mL of sodium dithionite solution was added to achieve a final concentration of 0.25–2.0 mg/mL. This dilution yielded a final tHb of 10.5–11.5 g/dL. For each prepared blood sample, a separate 0.3 mL draw was collected for measurement by a clinical-grade CO-oximeter (Avoximeter 7000, Accriva Diagnostics, Piscataway, NJ). We observed that while adding sodium dithionite powder directly to blood produced significant amounts of undesired methemoglobin within minutes (5–10%, measured by CO-oximetry), our approach generally resulted in very low methemoglobin levels for the duration of testing (0.0–0.8%). Ground truth CO-oximetry  $\text{SO}_2$  was generally within 2% of the intended target setpoint.

Phantoms were imaged using our previously described custom PAI system [36,37]. Briefly, this system combines a tunable optical parametric oscillator (Opotek, Inc., Carlsbad, CA) with a research-grade ultrasound acquisition system (Vantage 128, Verasonics, Inc., Kirkland, WA), using a 128-element linear array transducer with 7 MHz center frequency and 0.3 mm element spacing (L11-4 v, Verasonics). After placing the bulk phantom on the stage, a thin layer of water was added to the top surface, on which the skin layer was placed, and excess water was squeezed out by gentle compression. Additional water was placed on top of the skin layer for coupling to the array transducer. After installing the skin layer, ink or blood was injected into the tubes and imaging was performed. After imaging experiments in one skin layer were completed, the skin layer was removed and the next layer was installed as above. For each condition, 20 multispectral imaging scans were performed from 700–895 nm in 5 nm steps, with maximum radiant exposure of  $10.3 \pm 0.4 \text{ mJ/cm}^2$  at 750 nm (measured using a 3.5 mm diameter aperture).

Images were log-compressed and displayed at 60 dB dynamic range (80–140 dB signal amplitudes) to prevent image saturation, but fluence correction and spectral unmixing were performed on raw, pre-compressed images to preserve photoacoustic amplitude spectra. Fluence correction was performed using our previously described 1D diffusion dipole model (1D-DDM) algorithm [38], which considered an elliptical optical beam with radii of 1.75 cm and 0.25 cm that was offset 9 mm from the image plane (matching our PAI system's illumination geometry). We also used optical property spectra measured in PVCP batch control samples as inputs to the fluence model, which represents a best-case scenario for fluence correction.  $\text{SO}_2$  was calculated using image data from 725–850 nm and solving the following system of equations ( $P = EC$ ) for the relative concentrations of  $\text{HbO}_2$  ( $c_{\text{HbO}_2}$ ) and Hb ( $c_{\text{Hb}}$ ):

$$P = \begin{bmatrix} p(\lambda_1) \\ p(\lambda_2) \\ \vdots \\ p(\lambda_n) \end{bmatrix} = \begin{bmatrix} \varepsilon_{\text{HbO}_2}(\lambda_1) & \varepsilon_{\text{Hb}}(\lambda_1) \\ \varepsilon_{\text{HbO}_2}(\lambda_2) & \varepsilon_{\text{Hb}}(\lambda_2) \\ \vdots & \vdots \\ \varepsilon_{\text{HbO}_2}(\lambda_n) & \varepsilon_{\text{Hb}}(\lambda_n) \end{bmatrix} \begin{bmatrix} c_{\text{HbO}_2} \\ c_{\text{Hb}} \end{bmatrix} = EC$$

where  $p$  is fluence-corrected photoacoustic image amplitude at wavelength  $\lambda_i$ , and  $\varepsilon_{\text{HbO}_2}$  and  $\varepsilon_{\text{Hb}}$  are known extinction coefficients of  $\text{HbO}_2$  and Hb at wavelength  $\lambda_i$  [25].  $C$  was determined using the pseudo-inverse method [39], then used to calculate  $\text{SO}_2$ :

$$C = (E^T E)^{-1} E^T P, \text{SO}_2 = \frac{C_{\text{HbO}_2}}{C_{\text{HbO}_2} + C_{\text{Hb}}}$$

Maximum imaging depth was determined by calculating signal-to-background ratio (SBR), a common target contrast metric [21], for each tube, then interpolating the depth for which  $SBR = 2$ . Both SBR and target  $SO_2$  estimates were computed from images using 1.5 mm (horizontal) x 0.5 mm (vertical) regions of interest (ROIs) drawn over the top tube wall signal for each tube. All nonzero pixels within the ROI were averaged. Overall  $SO_2$  accuracy for each target depth and phantom skin layer was characterized using root-mean-square difference (RMSD), a metric used in ISO/IEC standards for pulse oximeter testing [20]:

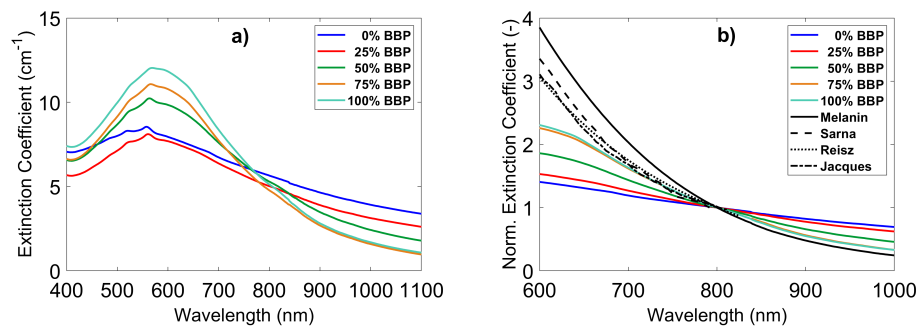
$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n (SO_{2,PA,i} - SO_{2,CO,i})^2}$$

where  $n$  is the number of measurement pairs,  $SO_{2,PA,i}$  is  $SO_2$  measured by PAI, and  $SO_{2,CO,i}$  is  $SO_2$  measured by CO-oximetry

### 3. Results

#### 3.1. Phantom characterization

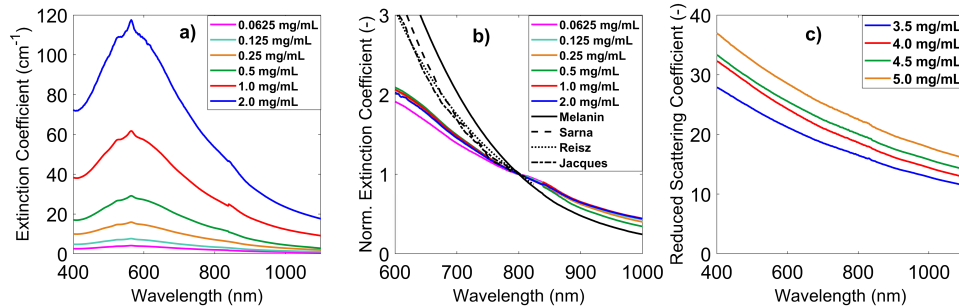
Alcohol-soluble nigrosin was generally compatible with PVC, and the addition of 2% ethanol to PVC had a negligible effect on the acoustic properties. Samples with 2% ethanol had no significant difference in speed of sound ( $1463 \pm 1.4$  m/s vs.  $1464 \pm 0.9$  m/s) and showed slightly higher acoustic attenuation of  $(0.411 \pm 0.002)f^{(1.70 \pm 0.01)}$  vs.  $(0.366 \pm 0.002)f^{(1.68 \pm 0.01)}$  in samples without ethanol. The optical extinction spectrum of nigrosin-doped PVC varied strongly with composition. Adding nigrosin as an ethanol solution (2% of TMM by volume) rather than directly to liquid plasticizers improved solubility and resulted in higher overall extinction as well as a much steeper slope of the extinction spectrum in the near-infrared range (data not shown). The extinction spectrum also varied significantly with BBP/DEHA plasticizer ratio, with increasing peak absorption and spectral slope from 600–1000 nm as BBP content increased (Fig. 2(a)). Results showed that PVC formulations containing 0–2 mg/mL nigrosin added as an ethanol solution to a 25% DEHA, 75% BBP plasticizer mixture achieved an extinction spectrum similar to that of melanin from 700–900 nm ( $10\% \pm 5.4\%$  error), with monotonically decreasing extinction with increasing wavelength (Fig. 2(b)). As this wavelength range overlaps with the range used by many PAI devices, this TMM formulation was selected for constructing phantoms used in imaging experiments.



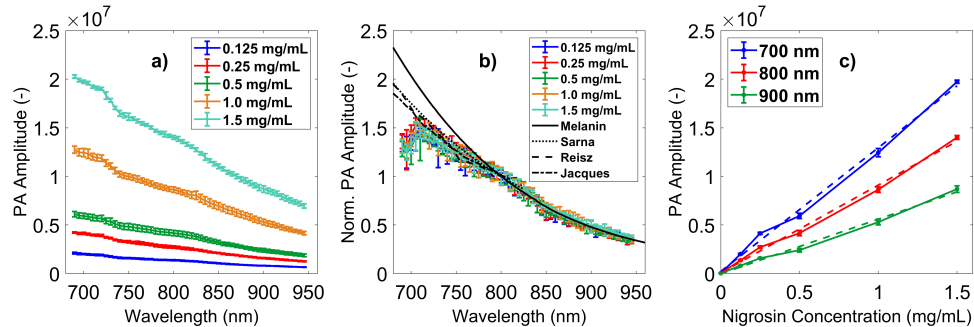
**Fig. 2.** a) Extinction spectra for various liquid plasticizer solutions with 0.25 mg/mL nigrosin. b) Extinction spectra normalized at 800 nm, compared to our data and literature data for melanin extinction [27–29].

PVCP samples doped with nigrosin introduced by ethanol solution were found to achieve high extinction coefficients beyond the range of target values (Fig. 3(a)), with a spectral slope

similar to that of melanin from 700–900 nm ( $18\% \pm 10\%$  error, Fig. 3(b)). PVCP samples of epidermis-mimicking formulations doped with  $\text{TiO}_2$  achieved high, epidermis-mimicking levels of scattering (Fig. 3(c)), and PVCP samples based on lower  $\text{TiO}_2$  concentrations used for the dermis and bulk phantoms achieved breast-like scattering of  $\sim 10 \text{ cm}^{-1}$  at 800 nm consistent with our previous published work [24] (data not shown). PAI of nigrosin-ethanol solutions injected into tubes in a water bath presented strong photoacoustic response, and photoacoustic spectra were highly consistent with both measured extinction spectra (Fig. 4(a)) and melanin spectra from 700–900 nm ( $7.9\% \pm 8.1\%$  error, Fig. 4(b)). Photoacoustic amplitude was also found to be highly linear vs. nigrosin concentration, with  $R^2$  values above 0.99 for 700, 800, and 900 nm (Fig. 4(c)). PVCP epidermis layer thickness was measured as  $62.2 \pm 7.6 \mu\text{m}$  based on OCT imaging of one sample. (Fig. S3).



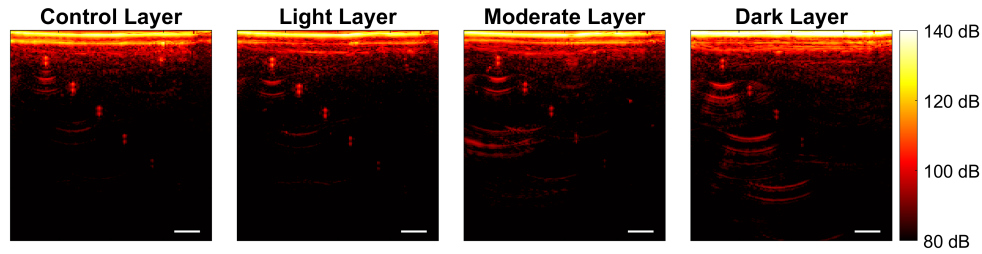
**Fig. 3.** a) Extinction spectra of PVCP vs. nigrosin concentration. b) Normalized extinction spectra compared to literature data for melanin [27–29]. c) Reduced scattering coefficient spectra of PVCP vs.  $\text{TiO}_2$  concentration.



**Fig. 4.** a) Photoacoustic image amplitude spectra for nigrosin-ethanol solutions in tubes within a water bath. b) Normalized photoacoustic spectra vs. our data and literature data for melanin extinction [27–29]. c) Linearity of photoacoustic image amplitude vs. nigrosin concentration at multiple wavelengths.

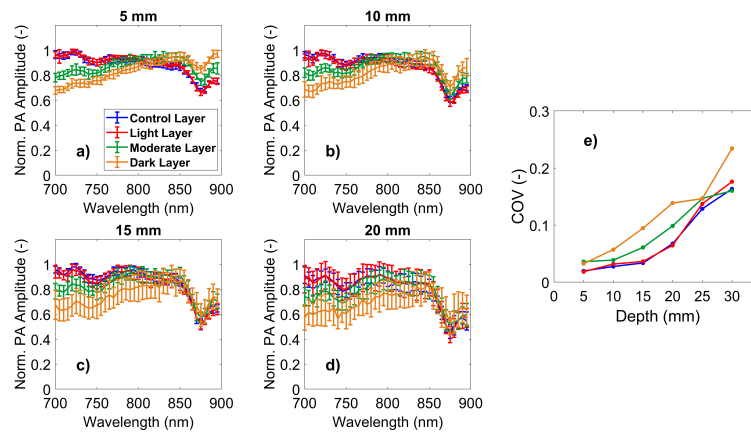
### 3.2. Image quality

As shown in Fig. 5, several differences in photoacoustic images were observed in phantoms with India ink-filled tubes. As phantom epidermal pigmentation level increased, target brightness generally decreased, with some deeper targets becoming nearly undetectable by visual inspection. Changes in target amplitude were difficult to visually identify due to log-compression; relative

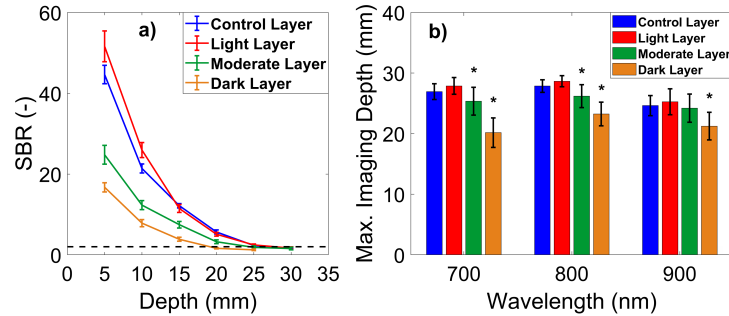


**Fig. 5.** PA images at 800 nm of the phantom filled with India ink for each of the four skin layers. Scale bar = 5 mm.

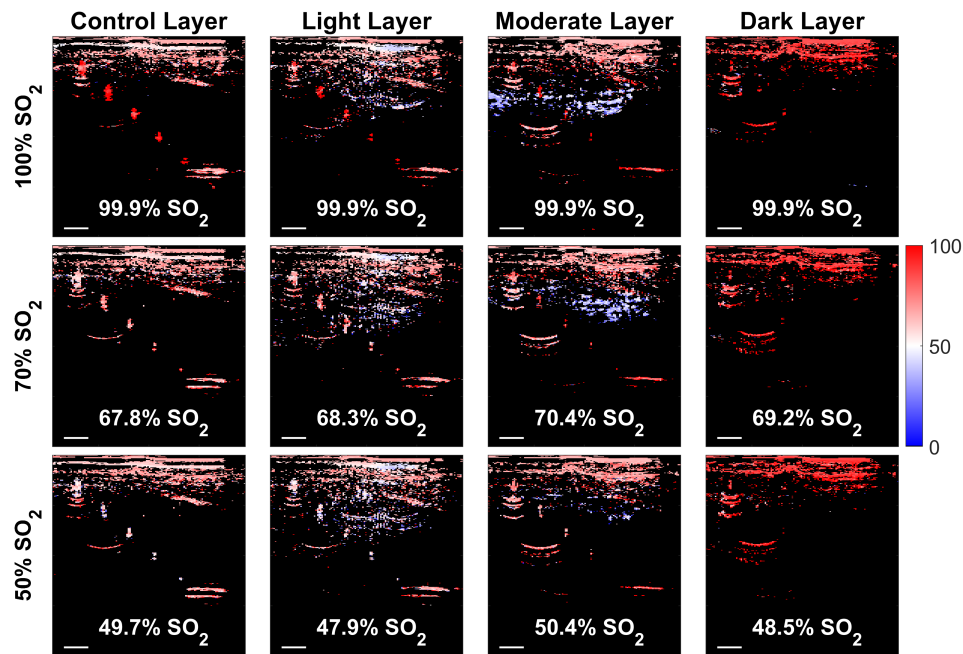
to the control layer, amplitude of the shallowest target slightly increased by  $0.6 \pm 0.2$  dB for the light layer, decreased by  $3.8 \pm 0.3$  dB for the moderate layer, and decreased by  $4.8 \pm 0.2$  dB for the dark layer. Reflection artifacts were observed beneath tubes as previously described [36], but with increasing signal amplitude as pigmentation level increased. Increasing pigmentation also caused stronger spectral coloring effects, especially for deeper targets (Fig. 6). Greater changes in normalized spectra were seen at shorter wavelengths, owing to higher absorption by nigrosin in that range. Also, signal quality decreased with increasing target depth and pigmentation level, as indicated by larger error bars in Fig. 6(a)-(d) and the coefficient of variation (COV) values in Fig. 6(e). Target SBR decreased at all depths (Fig. 7), and relative SBR reductions were similar across 5, 10, and 15 mm, with SBR decreasing to 55–61% and 32–37% of control layer SBR for the moderate and dark layers, respectively. Deeper targets showed smaller relative reductions as SBR approached the detection limit. Finally, while the light layer did not cause statistically significant decreases in maximum imaging depth, significant decreases were observed for the moderate and dark layers, with the latter resulting in  $\sim 25\%$  reduced imaging depth vs. the control layer at 700 nm, 800 nm, and 900 nm (Fig. 7(b)).



**Fig. 6.** Normalized PA spectra of ink-filled phantom tubes for tube depths of a) 5 mm, b) 10 mm, c) 15 mm, and d) 20 mm for each skin layer. e) Corresponding coefficient of variation (COV) values vs. target depth and skin layer.



**Fig. 7.** a) Target signal-to-background ratio (SBR) at 700 nm vs. depth for the four skin layers. Black dashed line indicates a threshold of  $SBR = 2$  chosen to define maximum imaging depth. b) Resultant maximum imaging depth for 700 nm, 800 nm, and 900 nm for each of the four skin layers. Asterisks denote statistically significant decreases in maximum imaging depth vs. control layer ( $p < 0.05$ ) using Student's t-test.

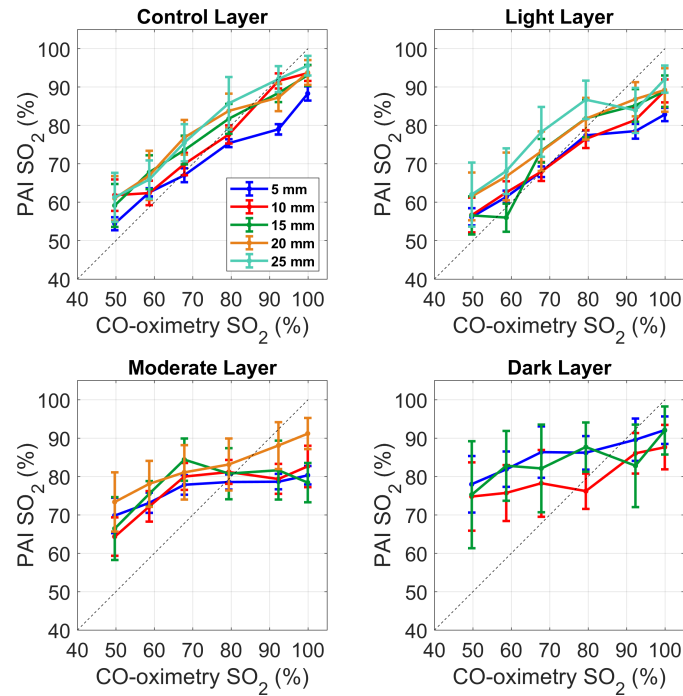


**Fig. 8.** PA oximetry maps of the phantom for each skin layer and three target  $SO_2$  levels (50%, 70%, and 100%). Ground-truth  $SO_2$  measurements by CO-oximetry are listed in each image. Scale bar = 5 mm. Black denotes that no value is displayed.

### 3.3. Oximetry measurement accuracy

$SO_2$  maps collected with each phantom layer showed several key differences in terms of overall image quality and appearance. As shown in Fig. 8, strong clutter signals were observed to overlap with target tubes and may have affected  $SO_2$  values. Some deeper targets did not show strong adjacent clutter, yet still presented  $SO_2$  bias. Reflection artifacts were observed beneath tubes as in ink phantom testing. Notably, superficial clutter for the dark layer possessed much higher  $SO_2$  values than for other pigmentation levels, and some deeper clutter features presented low



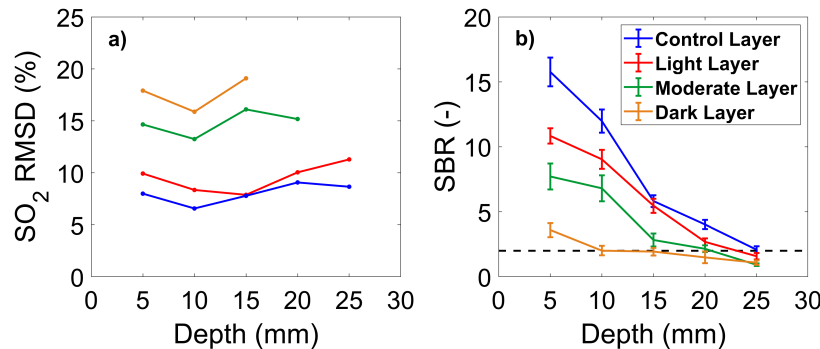


**Fig. 9.** PA SO<sub>2</sub> vs. CO-oximetry SO<sub>2</sub> for the control, light, moderate, and dark skin layer phantoms. The dashed black line denotes perfect agreement, and the error bars denote one standard error.

SO<sub>2</sub> (blue features in Fig. 8). To determine a best-case scenario for removing clutter from SO<sub>2</sub> maps, maps were also thresholded by only displaying SO<sub>2</sub> values within 1.5 mm (horizontal) x 2 mm (vertical) ROIs drawn over each target (Fig. S4). These idealized images used *a priori* knowledge of target locations to enable better overall visualization of targets, although some overlapping clutter is still observed near shallow targets. Both Fig. 8 and Fig. S4 show that maximum imaging depth and number of detectable targets decreased with pigmentation, with stronger differences than in India ink measurements. Different trends in SO<sub>2</sub> accuracy were also observed with increasing phantom pigmentation level (Fig. 9), with overall depth-averaged SO<sub>2</sub> RMSD of  $8.0\% \pm 1.0\%$ ,  $9.5\% \pm 1.4\%$ ,  $14.8\% \pm 1.2\%$ , and  $17.6\% \pm 1.6\%$ , for the control, light, moderate, and dark layers, respectively. The light layer showed similarly high accuracy as the control layer, but the moderate layer resulted in a flattened slope (reduced sensitivity to SO<sub>2</sub>) and positive SO<sub>2</sub> bias for lower SO<sub>2</sub> values. The dark layer resulted in even higher overall bias with a similar loss in sensitivity to SO<sub>2</sub> (reduced slope of the regression line). It is also evident that SO<sub>2</sub> measurement uncertainty increased with increasing pigmentation (Fig. 10), especially for deeper, targets with lower contrast. For the dark layer, targets below 15 mm were not visualized and did not yield consistent SO<sub>2</sub> measurements.

#### 4. Discussion

Characterization results indicated that nigrosin-doped PVCP mimics the extinction spectrum and photoacoustic response of melanin from 700–900 nm, and that this material can be fabricated into thin epidermis-mimicking layers bonded to dermis-mimicking layers. We did not specifically monitor phantom stability in this study, as PVCP was previously found to be stable for at least



**Fig. 10.** a) SO<sub>2</sub> RMSD values vs. target depth for the four skin layers. b) Target SBR at 800 nm vs. target depth for the four skin layers, with blood at 100% SO<sub>2</sub>.

6 months [24]. Nigrosin-PVCP skin phantoms qualitatively appeared to be stable for at least 1 month, but their long-term stability should be evaluated in future work. We did not explicitly evaluate photostability of nigrosin during PAI exposures, but photoacoustic spectra did not show drift over time during imaging (~6000 multispectral pulses per skin layer over all SO<sub>2</sub> cases). Nigrosin has been used in other PAI phantoms without reports of photostability concerns [40].

Image quality testing indicated that variable skin pigmentation can affect target detectability as evidenced by decreased contrast and maximum imaging depth. Compared to the control layer, maximum imaging depth was significantly reduced for the moderate and dark layers, but slightly increased for the light layer. This may be due to variable quality of acoustic coupling between the skin layer and bulk phantom. Phantoms with higher skin pigmentation also presented stronger image artifacts and clutter, which is attributed to stronger absorption and resultant photoacoustic signal generation within the epidermis. The dermis layer and bulk phantom were comprised of the same PVCP formulation to minimize acoustic reflections and image clutter. While skin and breast tissue actually have slight acoustic impedance mismatch of ~17% [41,42], tissue interfaces in humans are generally uneven and produce weaker specular reflections. Simple, planar phantoms have flat interfaces that can produce stronger, less biologically relevant acoustic reflections. Even after minimizing these effects, phantom testing showed that higher epidermal pigmentation can cause brighter image clutter artifacts, a trend that has also been reported by Mantri and Jokerst, who performed PAI of the dorsal forearm in human subjects with variable Fitzpatrick skin phototype [19]. Observed SO<sub>2</sub> bias may have been caused by both clutter overlapping with targets and by spectral coloring; since results indicated varying SO<sub>2</sub> bias with pigmentation level even for deeper targets, spectral coloring appears to significantly impact device accuracy. Additional clinical data is needed to establish expectations for typical photoacoustic clutter observed in patients and enable more precise replication of clutter in phantom testing.

Imaging depth in blood-filled phantoms was generally worse than in ink-filled phantoms, which may be because oximetry mapping excluded pixels with inadequate signal amplitude for at least one wavelength to avoid unmixing errors. Additionally, ink solutions simulated tHb of 13.5 g/dL, but variations in tHb of supplied blood and the need for sodium dithionite dilution reduced blood tHb to 10.5–11.5 g/dL, which should decrease photoacoustic signal amplitude by 15–22%. Superficial image clutter had higher apparent SO<sub>2</sub> for high pigmentation, which may be because the PAI system's laser beam spot was outside the image plane. Some clutter may be generated by light that diffuses through the skin layer and back to the transducer, undergoing spectral coloring and greatly reduced signal amplitude at lower wavelengths, which would result in positive SO<sub>2</sub> bias. However, the presence of clutter with very low SO<sub>2</sub> values suggests other clutter sources, such as out-of-plane signals generated beneath the laser spot (which would have high signal at

lower wavelengths due to nigrosin absorption and thus present lower  $SO_2$ ). We used *a priori* information to produce idealized  $SO_2$  maps with minimal clutter illustrating a best-case scenario, but more advanced approaches to detecting and eliminating clutter [43–45] may be needed to suppress signals not arising from blood vessels.

Phantom oximetry test results showed an overall decrease in  $SO_2$  accuracy with increasing skin pigmentation, with negative bias at high  $SO_2$  and positive bias at low  $SO_2$ . This result for high  $SO_2$  values is consistent with those reported by Mantri and Jokerst, who observed decreases in PAI measurements of radial artery  $SO_2$  (95–100% based on pulse oximetry) with increasing skin pigmentation in human subjects [19].  $SO_2$  maps in our study indicated that clutter thought to originate in the epidermal layer presented higher  $SO_2$  values, which is also consistent with Mantri and Jokerst's findings. However, their study only considered the radial artery in healthy subjects, which is expected to have a very high  $SO_2$  level. Our phantom-based test method offers an approach that can augment and complement clinical testing by enabling evaluation of device performance over a wider range of  $SO_2$  values than the range that can be safely produced in human subjects. This may be especially useful for evaluating PAI performance when imaging venous blood, as pulse oximetry is mostly insensitive to veins [46] and thus cannot provide a suitable noninvasive reference measurement for venous  $SO_2$ . This is also important because our phantom test results showed greater pigmentation-dependent  $SO_2$  bias at lower  $SO_2$  values. Finally, this study only evaluated pigmentation effects on a near-infrared PAI system from 700–900 nm; it is possible that due to its absorption spectrum, epidermal melanin may more strongly affect PAI devices that use shorter visible wavelengths [18] and have less impact on devices using longer wavelengths such as the second near-infrared window (1000–1700 nm) [47]. Notably, hyperspectral reflectance of human subjects has shown only small differences in skin reflectance from 1100–1800 nm [48].

In summary, phantom test results indicated that skin pigmentation can affect PAI device performance by attenuating signals, generating image artifacts, and causing spectral coloring effects that degrade image quality. These various mechanisms also contributed to reduced  $SO_2$  accuracy in PAI oximetry mapping with increasing skin pigmentation. While these results indicate a potential for pigmentation-dependent differences in device performance, further clinical evidence is needed to confirm and quantify these effects for different PAI device configurations. Furthermore, the clinical significance of these effects and potential impact on patient outcomes will depend considerably on a given device's intended use and clinical applications. The effects of melanin may be an especially important consideration towards developing PAI systems capable of absolute, quantitative photoacoustic oximetry measurements, where mitigation strategies, such as patient-specific spectral coloring correction algorithms [49,50], may be needed to ensure optimal performance across a diverse population.

## 5. Conclusions

We developed tissue phantoms that simulate variable epidermal melanin content and enable benchtop assessment of potential racial disparities in PAI device performance. Phantom testing indicated that variations in epidermal pigmentation level can significantly affect photoacoustic image quality and oximetry measurement accuracy. These phantoms represent useful regulatory science tools that enable identification of PAI device melanin susceptibility earlier in the device development process. These tools can also be used to evaluate new mitigation strategies and design optimizations intended to improve device robustness to skin pigmentation, thus helping to ensure PAI medical devices achieve equitable outcomes in all patients.

**Acknowledgments.** The authors thank Dr. Anant Agrawal from FDA/CDRH for use of his custom OCT system. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services. This article reflects the views of the authors and should not be construed to represent FDA's views or policies.

**Disclosures.** The authors have no conflicts to disclose.

**Data availability.** Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

**Supplemental document.** See [Supplement 1](#) for supporting content.

## References

1. M. W. Sjoding, R. P. Dickson, T. J. Iwashyna, S. E. Gay, and T. S. Valley, "Racial Bias in Pulse Oximetry Measurement," *N. Engl. J. Med.* **383**(25), 2477–2478 (2020).
2. J. R. Feiner, J. W. Severinghaus, and P. E. Bickler, "Dark Skin Decreases the Accuracy of Pulse Oximeters at Low Oxygen Saturation: The Effects of Oximeter Probe Type and Gender," *Anesth. Analg.* **105**(6), S18–S23 (2007).
3. P. E. Bickler, J. R. Feiner, and J. W. Severinghaus, "Effects of skin pigmentation on pulse oximeter accuracy at low saturation," *Anesthesiology* **102**(4), 715–719 (2005).
4. G. S. Barsh, "What controls variation in human skin color?" *PLoS Biol.* **1**(1), e27 (2003).
5. P. E. Bickler, J. R. Feiner, and M. D. Rollins, "Factors affecting the performance of 5 cerebral oximeters during hypoxia in healthy volunteers," *Anesth. Analg.* **117**(4), 813–823 (2013).
6. A. Afshari, R. B. Saager, D. Burgos, W. C. Vogt, J. T. Wang, G. Mendoza, S. Weininger, K. B. Sung, A. J. Durkin, and T. J. Pfefer, "Evaluation of the robustness of cerebral oximetry to variations in skin pigmentation using a tissue-simulating phantom," *Biomed. Opt. Express* **13**(5), 2909–2928 (2022).
7. P. Szabo, M. Wolf, H. Bucher, J. C. Fauchere, D. Haensse, and R. Arlettaz, "Detection of hyperbilirubinaemia in jaundiced full-term neonates by eye or by bilirubinometer?" *Eur. J. Pediatr.* **163**(12), 722–727 (2004).
8. S. Wainer, Y. Rabi, S. M. Parmar, D. Allegro, and M. Lyon, "Impact of skin tone on the performance of a transcutaneous jaundice meter," *Acta Paediatr.* **98**(12), 1909–1915 (2009).
9. J. Tsai, A. L. Chien, J. U. Kang, S. Leung, S. Kang, and L. A. Garza, "Hyperspectral measurement of skin reflectance detects differences in the visible and near-infrared regions according to race, gender and body site," *Acad Dermatol Venereol* **35**(5), E330–E333 (2021).
10. E. I. Neuschler, R. Butler, C. A. Young, L. D. Barke, M. L. Bertrand, M. Bohm-Velez, S. Destounis, P. Donlan, S. R. Grobmyer, J. Katzen, K. A. Kist, P. T. Lavin, E. V. Makariou, T. M. Parris, K. J. Schilling, F. L. Tucker, and B. E. Dogan, "A Pivotal Study of Optoacoustic Imaging to Diagnose Benign and Malignant Breast Masses: A New Evaluation Tool for Radiologists," *Radiology* **287**(2), 398–412 (2018).
11. L. Lin and L. H. V. Wang, "The emerging role of photoacoustic imaging in clinical oncology," *Nat. Rev. Clin. Oncol.* **19**(6), 365–384 (2022).
12. S. Manohar and M. Dantuma, "Current and Future Trends in Photoacoustic Breast Imaging," *Photoacoustics* **16**, 100134 (2019).
13. M. A. L. Bell, "Photoacoustic imaging for surgical guidance: Principles, applications, and outlook," *J. Appl. Phys.* **128**(6), 1 (2020).
14. K. Kratkiewicz, R. Manwar, A. Rajabi-Estarabadi, J. Fakhoury, J. Meiliute, S. Daveluy, D. Mehregan, and K. Avanaki, "Photoacoustic/Ultrasound/Optical Coherence Tomography Evaluation of Melanoma Lesion and Healthy Skin in a Swine Model," *Sensors* **19**(12), 2815 (2019).
15. Y. Zhou, S. V. Tripathi, I. Rosman, J. Ma, P. F. Hai, G. P. Linette, M. L. Council, R. C. Fields, L. H. V. Wang, and L. A. Cornelius, "Noninvasive Determination of Melanoma Depth using a Handheld Photoacoustic Probe," *J. Invest. Dermatol.* **137**(6), 1370–1372 (2017).
16. E. I. Galanzha, Y. A. Menyayev, A. C. Yadem, M. Sarimollaoglu, M. A. Juratli, D. A. Nedosekin, S. R. Foster, A. Jamshidi-Parsian, E. R. Siegel, I. Makhoul, L. F. Hutchins, J. Y. Suen, and V. P. Zharov, "In vivo liquid biopsy using Cytophone platform for photoacoustic detection of circulating tumor cells in patients with melanoma," *Sci. Transl. Med.* **11**(496), 1 (2019).
17. D. L. Longo, R. Stefania, S. Aime, and A. Oraevsky, "Melanin-Based Contrast Agents for Biomedical Optoacoustic Imaging and Theranostic Applications," *Int. J. Mol. Sci.* **18**(8), 1719 (2017).
18. X. T. Li, U. S. Dinish, J. Aguirre, R. Z. Bi, K. Dev, A. B. E. Attia, S. Nitkunanantharajah, Q. H. Lim, M. Schwarz, Y. W. Yew, S. T. G. Thng, V. Ntziachristos, and M. Olivo, "Optoacoustic mesoscopy analysis and quantitative estimation of specific imaging metrics in Fitzpatrick skin phototypes II to V," *J. Biophotonics* **12**(9), 1 (2019).
19. Y. Mantri and J. V. Jokerst, "Impact of skin tone on photoacoustic oximetry and tools to minimize bias," *Biomed. Opt. Express* **13**(2), 875–887 (2022).
20. IEC 80601-2-61:2017, "Medical electrical equipment — Part 2-61: Particular requirements for basic safety and essential performance of pulse oximeter equipment," <https://www.iso.org/standard/67963.html> (2017).
21. J. Palma-Chavez, T. J. Pfefer, A. Agrawal, J. V. Jokerst, and W. C. Vogt, "Review of consensus test methods in medical imaging and current practices in photoacoustic image quality assessment," *J. Biomed. Opt.* **26**(09), 1 (2021).
22. R. B. Saager, A. Quach, R. A. Rowland, M. L. Baldado, and A. J. Durkin, "Low-cost tissue simulating phantoms with adjustable wavelength-dependent scattering properties in the visible and infrared ranges," *J. Biomed. Opt.* **21**(6), 067001 (2016).
23. J. K. Tsou, J. Liu, A. I. Barakat, and M. F. Insana, "Role of ultrasonic shear rate estimation errors in assessing inflammatory response and vascular risk," *Ultrasound in Medicine and Biology* **34**(6), 963–972 (2008).

24. W. C. Vogt, C. Jia, K. A. Wear, B. S. Garra, and T. Joshua Pfefer, "Biologically relevant photoacoustic imaging phantoms with tunable optical and acoustic properties," *J. Biomed. Opt.* **21**(10), 101405 (2016).
25. S. A. Prael, M. J. van Gemert, and A. J. Welch, "Determining the optical properties of turbid media by using the adding-doubling method," *Appl. Opt.* **32**(4), 559–568 (1993).
26. A. Hariri, J. Palma-Chavez, K. A. Wear, T. J. Pfefer, J. V. Jokerst, and W. C. Vogt, "Polyacrylamide hydrogel phantoms for performance evaluation of multispectral photoacoustic imaging systems," *Photoacoustics* **22**, 100245 (2021).
27. T. Sarna and H. A. Swartz, "The Physical Properties of Melanins," in *The Pigmentary System: Physiology and Pathophysiology*, 2nd Edition, J. J. Nordlund, R. E. Boissy, V. J. Hearing, R. A. King, W. S. Oetting, and J.-P. Ortonne, eds. (Blackwell Publishing, Ltd., Oxford, UK, 2006).
28. J. Riesz, J. Gilmore, and P. Meredith, "Quantitative photoluminescence of broad band absorbing melanins: a procedure to correct for inner filter and re-absorption effects," *Spectrochim. Acta, Part A* **61**(9), 2153–2160 (2005).
29. S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**(11), R37–R61 (2013).
30. M. Cichorek, M. Wachulska, A. Stasiewicz, and A. Tyminska, "Skin melanocytes: biology and development," *Postępy Dermatol Alergol* **1**(1), 30–41 (2013).
31. P. Oltulu, B. Ince, N. Kokbudak, S. Findik, and F. Kilinc, "Measurement of epidermis, dermis, and total skin thicknesses from six different body regions with a new ethical histometric technique," *Turk J Plast Surg* **26**(2), 56–61 (2018).
32. S. Y. Huang, J. M. Boone, K. Yang, A. L. Kwan, and N. J. Packard, "The effect of skin thickness determined using breast CT on mammographic dosimetry," *Med Phys* **35**(4), 1199–1206 (2008).
33. J. J. Tronolone, J. Lam, A. Agrawal, and K. Sung, "Pumpless, modular, microphysiological systems enabling tunable perfusion for long-term cultivation of endothelialized lumens," *Biomed. Microdevices* **23**(2), 25 (2021).
34. S. L. Jacques, "Origins of Tissue Optical Properties in the UVA, Visible, and NIR Regions," in *Advances in Optical Imaging and Photon Migration 1996*, (Optical Society of America, Florida, 1996).
35. Y. Shimojo, T. Nishimura, H. Hazama, T. Ozawa, and K. Awazu, "Measurement of absorption and reduced scattering coefficients in Asian human epidermis, dermis, and subcutaneous fat tissues in the 400-to 1100-nm wavelength range for optical penetration depth and energy deposition analysis," *J. Biomed. Opt.* **25**(04), 1 (2020).
36. W. C. Vogt, C. Jia, K. A. Wear, B. S. Garra, and T. J. Pfefer, "Phantom-based image quality test methods for photoacoustic imaging systems," *J. Biomed. Opt.* **22**(09), 1–14 (2017).
37. W. C. Vogt, X. Zhou, R. Andriani, K. A. Wear, T. J. Pfefer, and B. S. Garra, "Photoacoustic oximetry imaging performance evaluation using dynamic blood flow phantoms with tunable oxygen saturation," *Biomed. Opt. Express* **10**(2), 449–464 (2019).
38. X. Zhou, N. Akhlaghi, K. A. Wear, B. S. Garra, T. J. Pfefer, and W. C. Vogt, "Evaluation of Fluence Correction Algorithms in Multispectral Photoacoustic Imaging," *Photoacoustics* **19**, 100181 (2020).
39. J. Laufer, C. Elwell, D. Delpy, and P. Beard, "In vitro measurements of absolute blood oxygen saturation using pulsed near-infrared photoacoustic spectroscopy: accuracy and resolution," *Phys. Med. Biol.* **50**(18), 4409–4428 (2005).
40. L. Hacker, J. Joseph, A. M. Ivory, M. O. Saed, B. Zeqiri, S. Rajagopal, and S. E. Bohndiek, "A Copolymer-in-Oil Tissue-Mimicking Material With Tuneable Acoustic and Optical Characteristics for Photoacoustic Imaging Phantoms," *IEEE Trans. Med. Imaging* **40**(12), 3593–3603 (2021).
41. K. Zell, J. I. Sperl, M. W. Vogel, R. Niessner, and C. Haisch, "Acoustical properties of selected tissue phantom materials for ultrasound imaging," *Phys. Med. Biol.* **52**(20), N475–N484 (2007).
42. F. A. Duck, *Physical Properties of Tissue* (Academic Press Inc, San Diego, CA, 1990).
43. M. Jaeger, J. C. Bamber, and M. Frenz, "Clutter elimination for deep clinical photoacoustic imaging using localised vibration tagging (LOVIT)," *Photoacoustics* **1**(2), 19–29 (2013).
44. M. A. Lediju Bell, N. Kuo, D. Y. Song, and E. M. Boctor, "Short-lag spatial coherence beamforming of photoacoustic images for enhanced visualization of prostate brachytherapy seeds," *Biomed. Opt. Express* **4**(10), 1964–1977 (2013).
45. M. Kuniyil Ajith Singh and W. Steenbergen, "Photoacoustic-guided focused ultrasound (PAFUSion) for identifying reflection artifacts in photoacoustic imaging," *Photoacoustics* **3**(4), 123–131 (2015).
46. P. D. Mannheim, "The light-tissue interaction of pulse oximetry," *Anesth. Analg.* **105**(6), S10–S17 (2007).
47. P. K. Upputuri and M. Pramanik, "Photoacoustic imaging in the second near-infrared window: a review," *J. Biomed. Opt.* **24**(04), 1 (2019).
48. M. J. Mendenhall, A. S. Nunez, and R. K. Martin, "Human skin detection in the visible and near infrared," *Appl. Opt.* **54**(35), 10559–10570 (2015).
49. B. Cox, J. G. Laufer, S. R. Arridge, and P. C. Beard, "Quantitative spectroscopic photoacoustic imaging: a review," *J. Biomed. Opt.* **17**(6), 061202 (2012).
50. S. Tzoumas, A. Nunes, I. Olefir, S. Stangl, P. Symvoulidis, S. Glasl, C. Bayer, G. Multhoff, and V. Ntziachristos, "Eigenspectra photoacoustic tomography achieves quantitative blood oxygenation imaging deep in tissues," *Nat. Commun.* **7**(1), 12121 (2016).